Exploring Performances of LSTM and GRU in Neural Computation: Insights and Comparative Analysis

José Martín Véliz Zambrano.
Tecnológico de Monterrey, Querétaro, México.

Abstract

This review paper provides a comparative analysis of Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), two pivotal methods employed in the realm of neural computation. Delving into the efficiencies, operational mechanisms, and applications of both methods, this paper elucidates their respective strengths and weaknesses, highlighting their utility based on task requirements. This review also identifies methodological inconsistencies, evaluates techniques, and reveals challenges faced by LSTM and GRU when dealing with sequences of varying complexity. Through a comprehensive overview of these two methods, this piece aims to inspire future directions in research aimed at improving current limitations and enhancing the performance of LSTM and GRU. The findings underscore the continued relevance of LSTM and GRU in addressing temporal information storage and retrieval issues in RNNs, therefore reinforcing their critical roles in the development of neural computation.

Introduction

In the realm of neural computation, Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) have been instrumental as revolutionary learning methods facilitating effective storage and recollection of information over time. Tackling various complexities such as storing precise, real, long-duration values, these two methods have paved the way for improved recurrent neural networks (RNNs). However, with the introduction of GRU as a streamlined version of LSTM, questions arise regarding their comparative efficiency and ideal domains for their application. This review aims to resolve these ambiguities, providing a synthesized view of the current knowledge, identifying inconsistencies, evaluating different methodological approaches, and suggesting future research directions.

Comparative Analysis

Introduced by Hochreiter (1997), LSTM presents a powerful solution to a crucial issue observed in neural computation—prolonged errors backflow and neglect of short-term memory in recurrent backpropagation. The fundamental innovation of LSTM lies in its gradient-based approach, which empowers the model's unique ability to store and recall crucial information over extended periods. This capacity for long-term memory allows LSTM to outperform traditional methods that often suffer from the swiftly decaying information in the backpropagation process.

Later, the inception of GRU offered an intriguing alternative to LSTM. As detailed by Zhang and Lipton (2019), GRU shares much of its structure with LSTM. However, it introduces an elegant complexity reduction by eliminating certain aspects of LSTM's gates while maintaining comparable performance. With the multiplicative gating mechanisms, GRU represents a streamlined version of recurrent neural networks, designed to maintain the time-series data processing power of LSTM but with less computational burden.

The unique characteristics of LSTM and GRU find different areas of application. For instance, LSTM, with its memory cells' ability to retain information over lengthy periods, has been highly effective in fields as varied as speech processing, non-Markovian control, and music composition (Hochreiter, 1997).

Conversely, GRU's capacity to capture both short-term and long-term dependencies in sequential data has shown promise in traffic flow prediction (Fu, Zhang, Li, 2016). Its customization zone of capturing dependencies lends itself to performance surpassing traditional linear models such as ARIMA that struggle with the non-linear and stochastic nature of traffic flow.

Inconsistencies and Challenges

While the advancements mediated by LSTM and GRU have revolutionized the realm of neural computation, they are not without challenges. Crucially, LSTM and GRU’s performance can deteriorate with increases in sequence complexity (Cahuantzi, Chen, Guttle, 2018). The methods' susceptibility to such complexity variations reveals an important inconsistency in their application, potentially limiting their utility in certain domains. Particularly, it has been observed that GRUs outperform LSTMs on sequences with lower complexity levels. However, when it comes to processing high-complexity sequences, LSTMs have been found to exhibit greater efficiency.

Deep Learning and Neural Networks

As LSTM and GRU continue to forge new paths in neural computation, stimulating exploratory research is crucial to deepen our understanding of these methods. A promising avenue for future research could be to investigate the implementation of variably timed inputs and the incorporation of a higher number of hidden states in recurrent neural networks, known to amplify model performance. Expanding the scope of applications for LSTM and GRU to different domains and a wider variety of datasets can provide invaluable insights into the potential and limitations of these advanced neural network frameworks. Further, it is pertinent to explore enhancements or modifications that tactfully navigate the limitations inherent in LSTM and GRU when operating on divergent complexities and lengths of sequential data.

Conclusion

Although LSTM and GRU both aim to enhance memory retention in recurrent networks, they diverge in their operational mechanisms and their respective areas of application. Each presents unique strengths and weaknesses, and their selection often hinges upon the specific task requirements. Continued research should focus on navigating the limitations inherent to each of these network structures, and on attempts to extend their effectiveness in diverse application domains. This includes developing adjustments or additional methods that can work in synergy to overcome current challenges or further enhance LSTM and GRU's performance. Undoubtedly, LSTM and GRU will continue to be focal points of research in neural computation due to their critical roles in addressing the prevalent challenges of storing and retrieving temporal information.

6. References

[1] Hochreiter, S. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735-1780.

[2] Zhang, A., & Lipton, Z. C. (2019). Gated Recurrent Units (GRU). Dive into Deep Learning. Amazon and Carnegie Mellon University.

[3] Fu, R., Zhang, Z., & Li, L. (2016). Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction. Tsinghua University.

[4] Phi, M. (2018). Illustrated Guide to LSTM's and GRU's: A Step by Step Explanation. Towards Data Science.

[5] Cahuantzi, R., Chen, X., & Guttle, S. (2018). A Comparison of LSTM and GRU Networks for Learning Symbolic Sequences. The University of Manchester, Department of Mathematics.